

Message from Director, Office of Portfolio Analysis



George Santangelo, Ph.D.
Director

The biomedical research enterprise managed by Institutes, Centers, and Offices (ICOs) of the National Institutes of Health (NIH) spans a complex landscape of topics as diverse as epidemiology, social psychology, tissue engineering, and physical chemistry. Optimizing management of the **more than 38,000 NIH-funded research project grants** each year (totaling more than \$21.5B in FY 2019) requires a deep understanding of that inherent complexity. Toward that end, the Office of Portfolio Analysis (OPA) will continue to support data-driven decision making by developing and disseminating validated, carefully normalized approaches that can analyze past decision making, measure the resulting impact, and forecast the productivity of NIH portfolios, including their potential for successful bench-to-bedside translation. We will also continue to pursue and develop discoveries in metascience that help to promote rigorous, innovative, and

transformative research. These components are central to our ongoing commitment to support good stewardship of America's investments in the effort to obtain fundamental knowledge about the nature and behavior of living systems and apply that knowledge to enhance health, lengthen life, and reduce illness and disability.

Over the past several years, OPA has developed artificial intelligence and machine learning (AI/ML) approaches that use a broad spectrum of data sources to create metrics that standardize quantitation of the overall productivity of research investments, including their clinical and technological impact. These data sources include a variety of document types (e.g., grant applications and awards, publications, patents, and clinical trials); associated metadata (e.g., affiliations, authorship, and citations); natural language processing (NLP)-derived elements; libraries of drugs and lead compounds; electronic health records; and economic data. Our validation process ensures that OPA metrics are clearly defined, carefully formulated, scientifically rigorous, and fully transparent through public access to all raw data and peer review of the underlying science. Current OPA tools—which include *iCite*, *iSearch*, and most recently, the **COVID-19 Portfolio**—also apply AI/ML, NLP, and other analytical schema that provide decision makers with sophisticated ways to search documents, associated metrics and metadata, and other linked datasets. Although these metrics and tools have been shown to improve decision making, they are designed to augment, not substitute for, human judgment.

Our other metascience research projects, which leverage major advances in supercomputing architecture and software design, are rapidly expanding our ability to improve data quality and provide valuable insights. For example, OPA is optimizing our AI/ML disambiguation algorithms to create the comprehensive, interoperable databases needed to fully capture the impact of biomedical research and the scientists who create it. We are also using AI/ML to achieve a degree of dimensionality reduction that allows comparison of the pre-decisional portfolios of private and public funders, facilitating identification of potentially overlapping investments while honoring intellectual property rights and all other data-sharing restrictions. Finally, our efforts to build AI/ML that detects transformative emerging areas of research leverage citation analytics, graph theory, and NLP in a novel, consistent approach.

The three objectives in this plan focus on the OPA mission to continue the development and refinement of analytics and tools that can help manage portfolios, identify high-performing areas of research, and detect exciting new scientific opportunities. We will also maintain our robust training, service, and coordination efforts to improve data driven decision making across NIH and beyond, for example, as other federal agencies continue to seek our assistance. I look forward to this next phase of development of OPA as we work together with colleagues inside and outside of NIH in pursuit of these important goals.

George M. Santangelo, Ph.D.
Director, Office of Portfolio Analysis